DOCUMENT RESUME

ED 092 880 CS 001 139

AUTHOR Hood, Joyce; Kendall, Janet Ross
TITLE A Qualitative Analysis of Oral Reading Miscues of Reflective and Impulsive Second Graders: A Follow-up Study.
PUB DATE Apr 74
NOTE 30p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, April 15-19, 1974)

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS Grade 2; *Miscue Analysis; Oral Reading; Reading; Reading Ability; Reading Comprehension; *Reading Processes; *Reading Research; Reading Skills; *Reading Speed

ABSTRACT

Subjects within one reading level only were included in this study to determine reflective and impulsive subject groups on the basis of Matching Familiar Figures (MFF) test time and error scores within that reading level rather than over all levels, to use the same reading test passages for all subjects, and then to apply statistical tests of significance to any observed differences between subject groups. The subjects were all the second graders in the same midwestern city whose teachers had chosen a certain grade-level reader for their reading instruction. The MFF test was administered to all of the subjects. The subjects were then asked to read test passages aloud and an examiner audio-tape recorded and noted errors in the children's oral reading. The miscues were coded for: word order changes, substitutions of meaningful words, insertions, omissions, skipped words, and punctuation. Each miscue was categorized as graphically similar or dissimilar, as corrected or uncorrected, and as to its contextual appropriateness. The results indicated that no significant differences were found between reflective (REF) and impulsive (IMP) subjects in the proportions of graphically similar and contextually appropriate miscues, and IMP children appeared to correct fewer unacceptable miscues than REF subjects. (WR)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

14.07

ED 092880

A Qualitative Analysis of Oral Reading Miscues of Reflective and Impulsive Second Graders: A Follow-up study

Joyce Hood and Janet Ross Kendall

The University of Iowa

Qualitative analysis of oral reading miscues is a relatively recent innovation in reading research. In contrast to quantative analysis, which simply determines the number of miscues (e.g. substitutions, omissions, and insertions) that are made, qualitative analysis provides an evaluation of the miscues that are counted. Each miscue is evaluated for the similarity of itr printed form to the graphic characteristics of the text, its semantic and syntactic appropriateness to the context, and its status as corrected or uncorrected by the reader. In the following paragraph a qualitative analysis of three miscues will be explained to illustrate the procedure.

In the sentence Johnny was happy with his birthday gift, if the reader says present instead of gift his substitution is semantically and syntactically appropriate since a present is the same thing as a gift, but its graphic form is different from that of the text word except for the final t. If the

CS 001139

reader says goat instead of gift, his substitution is not semantically appropriate but it is syntactically appropriate since it is a noun. The graphic form of this substitution has three things in common with the text word: its length, the initial g, and the final t. If the reader says brother instead of birthday, his substitution is not semantically appropriate but the two words are somewhat similar in graphic form. This substitution is syntactically appropriate if only the preceding portion of the sentence is considered (Johnny was happy with his brother), but it is not syntactically appropriate in the sentence as a whole (Johnny was happy with his brother gift).

The graphic similarity and the semantic and syntactic appropriateness of each of a reader's miscues are evaluated in relation to the miscue's status as corrected or uncorrected in order to characterize the reader's oral reading behavior. In the preceding examples the substitution present/gift is considered a better miscue qualitatively than goat/gift or brother/birthday. Although the latter two miscues are more graphically similar to the text, they do not approximate the meaning represented in the text. The more efficient reader would be expected to correct the miscues goat/gift and brother/birthday but to leave present/gift uncorrected.

Qualitative miscue analysis has been employed to study the relative contributions of graphic cues and contextual (semantic and syntactic) cues in the reading process. For example, studies of college students reading material in which letters were rotated or reversed, and bilingual readers reading mixed bilingual passages, have suggested that these adult readers are more sensitive to the contextual relationships of the words they are reading than to the graphic characteristics of the text words (Kolers, 1969). Investigations of the oral reading of first-graders have yielded similar results (Clay, 1969, Weber, 1970). For both first-graders and adult readers a higher proportion of their miscues were contextually appropriate than were graphically similar to the text.

Differences in correction rates have further illuminated the relative importance of graphic and contextual cues. Both Kolers and Weber reported that their subjects tended to correct oral reading miscues that were inappropriate to the context and to leave uncorrected the contextually appropriate miscues even when the printed forms of the miscues were graphically dissimilar to the text words.

Qualitative miscue analysis has demonstrated some differences in the oral reading behavior of high and low achievers. Although Weber found no differences between ability groups in the use of contextual constraints she did find that high-achieving first-graders seemed to be more sensitive than low achievers to graphic constraints. Biemiller (1970) found differences between high and low achieving first-graders in the use of graphic cues even when they were compared during equivalent stages in the development of reading ability. Both Clay (1969) and Weber (1970) reported greater proportions of corrected miscues for high-achieving than for low-achieving readers.

Recently there has been a growing interest in the possibility that children's educational development may be enhanced by modifying instructional programs to suit individual differences in cognitive style. For example, it is possible that differences in the oral reading behavior of high and low achievers may be related to cognitive style, and if so, that instructional approaches will differ in effectiveness depending upon the cognitive styles of the children who receive the instruction.

One dimension of cognitive style which seems to relate to oral reading behavior is that of cognitive tempo. A reader who

typically corrects his miscues seems to be behaving reflectively according to Kagan's characterization of conceptual tempos (Kagan, Rosman, Day, Albert, and Phillips, 1964). That is, a reflective child pauses to evaluate the quality of his thinking and the accuracy of his conclusions. On the other hand, a reader who is not likely to correct his miscues seems to be behaving impulsively, That is, he reports the first idea he thinks of with minimal consideration of its appropriateness or quality.

According to Frank Smith" . . .fluent reading, and learning to read fluently, require a willingness to 'make mistakes'. And the extent to which a child is prepared to risk mistakes is directly related to the tolerance of the teacher in accepting them. (1971, p. 230)." Smith also emphasizes the importance of feedback. "Every response a child makes, every identification of a letter or word or meaning, should be given the feedback of whether it is right or wrong." (p. 229) A child's awareness of congruence between his response and the graphic and contextual cues represented in the text constitutes positive feedback; incongruity signals that another response should be tried, i.e., that a correction is needed. If the child is willing to risk mistakes but is not sensitive to incongruity as a signal to

correct, he may need additional feedback from a teacher. If he is not willing to hazard any guesses, however, he will have no opportunity for informative feedback and thus no opportunity to discover new rules for identifying words. A reflective child may be overly concerned with errors and may be denying himself the opportunity to test new rules. An impulsive child may not be concerned about the correctness of his responses and may be losing information about the efficacy of his rules.

There have been numerous studies of relationships between various measures of reading ability and various aspects of cognitive style, but very few studies have specifically dealt with oral reading behavior and its relationship to reflection-impulsivity. Kagan (1965) was the first to conduct such a study. He recorded the errors made by second-graders while reading aloud. He reported positive correlations between an index of reflection-impulsivity and the total number of oral reading errors, the number of graphically similar errors, and the error variable that combined meaningful and nonmeaningful substitution errors and suffix errors.

Because Kagan employed a quantitative rather than a qualitative analysis of miscues he did not determine the proportion

of the total number of errors that were graphically similar and the proportion of the errors that were meaningful substitutions. If he had done so, it would have been possible to learn something about the relative use of these two cue sources by reflective and impulsive children. For example, he could have examined the proportions of corrections in these two classes of errors by the two subject groups to investigate whether reflective children might be correcting too many meaningful errors or whether impulsive children might be correcting too few of their nonmeaningful ones.

One of the problems encountered in planning a qualitative analysis of the oral reading errors of reflective and impulsive children is the conflict between the procedure for formulating reflective and impulsive subject groups and the choice of appropriate selections to be read aloud. This problem is explained in the following paragraphs.

A subject's identity as reflective or impulsive is determined by the median score on the time required to make the first response to each item and total number of erroneous responses on Kagan's Matching Familiar Figures test (MFF) for the group of subjects currently being tested. (Kagan et. al., 1964). It seems reasonable to expect that the subjects who take the test

should be randomly selected, or, if an intact group of subjects is used, that the group should include a fairly normal range of subjects so that the median time and error scores for children of a given chronological age will be somewhat similar from study to study.

If a normal range of reading abilities is represented in the subject groups a problem arises in the selection of a reading test passage, however. If the passage is easy enough so that the poorest readers can read it without frustration it will be so easy that the best readers will produce no miscues for the analysis. This would present no problem in a quantitative analysis of oral reading miscues, but there is some concern that the quality of an individual's oral reading may differ depending upon the number of errors made (McCracken, 1967, Weber, 1968). If the test passage is difficult enough so that the best readers produce some errors it may be so difficult for the poorest readers that a qualitative analysis will not reveal their typical oral reading behavior. Further, in Kagan's discussion of the minimal relationship he found between word errors and impulsivity for low verbal boys, he commented that tendencies toward reflection present no advantage in a situation where a subject has no

alternative hypotheses on which to reflect. He advised, "A preferred disposition for reflection or impulsivity is maximally influential at (the) intermediate level of mastery (1965, p. 626)."

An attempt to provide test passages of intermediate difficulty for all subjects was employed in a study by Hood, Kendall and Roettger (1973). They obtained samples of the oral reading of all the 79 first-graders enrolled in one elementary school in a midwestern city (population 45,000) by asking each child to read orally the next story in his reader following the one currently studied in his reading group. Only 50 subjects were classifiable as reflective or impulsive*. Of these, five reflective and five impulsive children read stories from the first reader, twenty reflective and seventeen impulsive children read stories from primers, and three children, all of them impulsive, read stories from preprimers. The total number of errors for each subject ranged from 0 to 6 for reflectives and from 0 to 15 for

---

*The remaining 49 subjects included those who were either above or below the median MFF scores for the group on both response time and errors and therefore could not be classified as reflective or impulsive, and one impulsive child who made 78 errors in his reading of a 100-word selection.

impulsives, indicating that the stories were within the intermediate mastery level for all subjects, though undoubtedly too easy for some.

A qualitative analysis of the oral reading miscues in the Hood et al., study revealed that reflective children made more repetitions, regressions, and proportionately more corrections of miscues than impulsive children but impulsive children made more miscues. There appeared to be no substantial differences between the proportions of contextually appropriate and graphically similar miscues for the two groups but there were differences in the proportions of corrected miscues in some of the miscue classes. The reflective children corrected proportionately more semantically inappropriate miscues, more miscues that were syntactically inappropriate to the following portion of the sentence, and more graphically dissimilar miscues than the impulsive children.

Obtaining samples of oral reading from a variety of stories had restricted the range of errors in the Hood et al. study, but because each of the subjects did not read the same test selection statistical tests of the significance of the differences between subject groups could not be employed. The observed

differences between reflective and impulsive children suggested that perhaps these impulsive children needed training in self-correction. Further, because there were no differences in the correction of miscues that were appropriate in the context, the results did not suggest that the reflective children were overly concerned with errors.

The Hood et al. study revealed nearly equal distributions of reflective and impulsive readers within reading levels. Therefore it seemed appropriate to plan a study which included subjects within one reading level only, to determine reflective and impulsive subject groups on the basis of MFF time and error scores within that reading level rather than over all levels, to use the same reading test passages for all subjects, and then to apply statistical tests of significance to any observed differences between subject groups. This is the plan of the present study.

Method

Subjects The subjects were all the second-graders in the same midwestern city as in Hood et al. whose teachers had chosen a certain grade-level reader for their reading instruction. Subjects who had participated in the previous study were excluded.

There were 166 subjects for whom MFF scores and oral reading samples were obtained. The median time to the first response on the MFF was 9.03 seconds, the median number of errors was 12.5. Extreme groups of 25 reflective (REF) and 25 impulsive (IMP) subjects were chosen. The REF subjects had time scores at or above 12.93 and error scores at or below 9. The IMP subjects had time scores at or below 7.7 and error scores at or above 16.

Testing Procedure The MFF test was administered by the second author and the oral reading samples were obtained by the first author who did not know the subjects' MFF time and error scores. Two test passages were employed, both 233 words long, one of high-second grade readability level and the other of high fourth-grade level. The subjects were told that the examiner was interested in learning what kinds of stories were best for second-graders to read. The subjects were asked to read the stories aloud just as they always read, and, if they came to an unknown word, to try to sound it out or guess what it was or skip it. The examiner gave no help with unknown words. The children's oral reading was audio-tape recorded. The examiner also made a written record of errors on a copy of the test passage while each child was reading.

Training of Scorers The scorers were five female college graduates, two of whom had been teachers of primary level reading. The oral reading samples of five subjects were used in training the scorers. Each subject was assigned to a scorer who scored that subject's errors and then scored the errors of the remaining four subjects. Each scorer compared her scoring of the subject assigned to her with that of every other scorer and discussed scoring differences with each scorer individually until the differences were resolved, making changes on the scoring sheet to reflect the changes agreed upon. The revised scores from the first scorer became the subject scores used in data analyses for these five subjects.

Scoring Procedure Each of the 45 oral reading samples which had not been used during the training period was scored independently by the five scorers who did not know the subjects' identities as REF or IMP. Each scorer listened to the audio-tape recording, corrected a photocopy of the examiner's written record, and then recorded the errors on a coding sheet. Five groups of nine subjects were formed. The sequence of scoring was balanced among scorers so that each group was scored once in each position in the sequence.

The miscues which were coded were: word order changes, substitutions of meaningful words, nonsense substitutions (such as strample/stumble), insertions, omissions (assumed to be inadvertent omissions), skipped words (assumed to be unknown words), and punctuation miscues. Each miscue was categorized 1) as graphically similar or dissimilar (GS or NGS) depending on whether the first letter of its printed form was the same as that of the text word, (insertions, omissions, skipped words, and punctuation miscues are always (NGS)2) as corrected or uncorrected by the subject, (COR, or NOT), and 3) as to its contextual appropriateness. The context in which each miscue was evaluated for appropriateness included the subject's previous uncorrected miscues but did not include any of the subsequent miscues. The miscue might be evaluated as not contextually appropriate (NCA), appropriate when only the preceding but not the following portion of the sentence was considered (PRE), appropriate to the sentence in which it occurred but not carrying the author's intended meaning (SEN), or appropriate in the context of the whole passage, i.e. consistent with the author's intended meaning (PASS). A

miscue was considered contextually appropriate only if it was both syntactically and semantically appropriate.*

The scores which were derived from the coding of errors were: total number of errors (MISCUES), proportions of graphically similar and dissimilar miscues (GS and NGS), proportions of miscues contextually appropriate (PASS) and inappropriate (NCA, PRE, and SEN), proportions of corrections (CORRECTION), and proportions of corrections within categories of graphic similarity and contextual appropriateness (GS-COR, NGS-COR, NCA-COR, PRE-COR, SEN-COR, and PASS-COR).

Total errors and proportions for each of these 45 subjects were determined for each scorer on the separate stories and on the stories combined. The scores for each subject were the mean of the total error scores and the means of the proportions over the five scorers. In situations where fewer than three scorers had recorded miscues for a subject within a given miscue category, the mean proportion of corrections was not determined. If three or more scorers had recorded miscues the mean proportion of

---

*A copy of the coding instructions and a paper explaining the rationale of the coding procedure are available from the first author.

corrections was determined for that subject over the number of scorers who had recorded miscues. Statistical analyses were planned to test the significance of differences between REF and IMP subjects for all of the subject scores for the stories combined. Because thirteen related t-tests were to be performed, the alpha level for the t-tests was set at .05/13, or .0038, as recommended by Dunn (1961).

Results

Table 1 presents the means, ranges, and standard deviations of miscues made by REF and IMP subjects on the separate stories and on the stories combined. The means were lower for REF than for IMP subjects, but the difference between the means on the combined stories was not statistically significant ($t=1.81$, $p < .10$). There was a wide range in the number of miscues made by subjects in the two groups. When the mean number of MISCUES was determined over REF and IMP subjects combined, and each subject was categorized as falling above or below the mean, it was determined that a significantly greater number of REF subjects fell below the mean (REF, 21, IMP, 12, corrected chi-square 5.70, $p < .02$).

The mean proportions of miscues in the various categories for REF and IMP subjects are presented in Table 2. It may be seen that

the differences between subject groups were consistent for the separate stories and for the stories combined. The reliabilities of the mean proportions for the stories combined are included in the tables. These reliabilities are estimates of the correlations which would be expected between these proportions and the proportions which would be obtained if the oral reading samples were scored by a second set of five scorers. The reliabilities are based on an analysis of variance among scorers according to the procedure described by Lindquist (1953).*

No significant differences were found between REF and IMP subjects in the proportions of miscues classified as NCA, PRE, SEN, or PASS, nor in the proportions of GS miscues. Because there were proportionately fewer miscues categorized as contextually appropriate at the sentence but not the passage level (SEN) and because the reliability of this categorization was relatively low, the SEN and PASS categories were combined and the proportions of miscues in the combined categories were determined for each subject. The mean proportions of miscues in the SENPASS category did not differ significantly for REF and IMP subjects.

---

*The reliability of the total error scores is .99.

REF subjects corrected proportionately more miscues than IMP subjects as shown in Table 2. Table 3 presents the mean proportions of corrections within the various miscue categories. Because there were so few miscues categorized as SEN and PASS, the proportions of corrections are reported only for the combined category (SENPASS). In several instances there were subjects who had no miscues in a given category; subjects with no miscues were not included in the mean proportions of corrections within categories. In Table 3 it may be seen that the proportions of corrections differed significantly only within the categories of graphically similar miscues (NGS) and miscues contextually appropriate in the preceding but not the following context (PRE).

Discussion

There were more REF than IMP subjects with low MISCUE scores, and the mean number of MISCUES for all REF subjects was lower than for IMP subjects, but the difference between REF and IMP means was not statistically significant. Thus these REF and IMP subjects performed somewhat like Kagan's subjects, but the difference between the subject groups in this study seems less pronounced. A number of procedural differences between the two studies should be noted.

Two procedural differences have to do with the selection of subjects and oral reading test passages: 1) Kagan's subjects were unselected as to reading ability. In the present study only subjects who were members of the grade-level instructional groups in each of the participating schools were included. 2) Kagan's subjects' most frequent error was failure to recognize a word due to lack of knowledge, which suggests that his test passages were relatively difficult for his subjects. In the present study fewer than 10 percent of the MISCUES were failures to respond and the most frequent error was word substitutions.

The intent of the present study was to restrict the range in the number of MISCUES made so that REF and IMP subjects could be more fairly compared with regard to the quality of their MISCUES. Even though the subjects were selected from only one instructional level, however, the MISCUE scores still varied widely. Hood et al. had found four IMP and no REF subjects in the lowest reading group and had excluded one of these IMP subjects because of an excessive number of errors. Thus, if the subjects in the present study had been unselected as to reading ability, the difference between REF and IMP means would probably have been greater, but the variance in MISCUE scores would probably have been much greater, also.

Two additional procedural differences between this and Kagan's studies have to do with the method of determining the relationship between reflection-impulsivity and the number of MISCUES made in oral reading: 1) Kagan reported correlations between indices of reflection-impulsivity and MISCUES for all his subjects. Mean numbers of MISCUES were reported for extreme groups of reflective and impulsive subjects in this study. 2) Kagan reported significant correlations between MISCUES and MFF time scores for boys and between MISCUES and MFF error scores for girls, but he did not report correlations of MISCUES with MFF time and MFF error scores in combination. In the present study both MFF time and MFF error scores were considered in forming subject groups, and boys and girls were not considered separately.

It should be emphasized that Kagan found significant correlations between MISCUES and MFF time scores only for boys and between MISCUES and MFF error scores only for girls. Whether he might have found a significant difference between REF and IMP means for extreme groups of boys and girls together cannot be determined. Out of 166 oral reading samples obtained for the present study, only the 50 samples for the subjects in the extreme REF and IMP groups have been scored. The scoring of the

samples for the remaining 116 subjects is planned. This will make it possible to determine the correlations between MISCUES and MFF time and error scores for boys and girls for comparison with Kagan's results.

The results of the qualitative analyses of miscues in this study are consistent with the results reported by Hood et al. No differences were found between REF and IMP subjects in the proportions of graphically similar and contextually appropriate miscues, but differences were found in the proportion of corrections in some categories of miscues. As in the Hood et al. study, IMP subjects corrected proportionately fewer miscues that were graphically dissimilar to the text and proportionately fewer miscues that were appropriate to the preceding but not the following context. Thus IMP children appear to correct fewer unacceptable miscues than REF subjects. This suggests that they may benefit from training in self-correction such as that described by Hood (1974).

Some categories of miscues are considered more acceptable than others. For example, miscues which are contextually appropriate at the sentence and/or passage level (SENPASS) may be near enough to the author's intended meaning not to require correction

by the reader. If REF subjects correct proportionately more miscues than IMP subjects, it has been argued, they may be correcting many acceptable miscues unnecessarily. Although there was an observed difference between REF and IMP subjects in the correction of SENPASS miscues in this study, the difference was not statistically significant. Further, neither REF nor IMP subjects appeared to correct SENPASS miscues as frequently as they corrected miscues in other categories. Therefore the results of the present study provide no evidence that these REF subjects are overly concerned with errors.

Biemiller, A. The development of the use of graphic and contextual information as children learn to read. Reading Research Quarterly, Fall 1970, 6, 75-96.

Clay, Marie M. Reading errors and self-correction behavior. British Journal of Educational Psychology, February 1969, 39, 47-56.

Dunn, O. J. Multiple comparisons among means. Journal of the American Statistical Association, 1961, 56, 52-64.

Hood, J. Why we burned our basic sight vocabulary cards. The Reading Teacher, 1974, 27 (6), 579-582.

Hood, Joyce, Kendall, Janet R., and Roettger, Doris M. An analysis of oral reading errors of reflective and impulsive first-graders. Paper presented at the AERA Convention, New Orleans, 1973.

Kagan, Jerome, Bernice Rosman, Deborah Day, Joseph Albert, and William Phillips. Information processing in the child: significance of analytic and reflective attitudes, Psychological Monographs, 1964, 78, 1-37.

Kolers, P. A. Reading is only incidentally visual. In K. S. Goodman & J. T. Fleming (Eds.) Diagnostic Viewpoints in Reading. International Reading Association, 1971, 121-133.

Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin Company, 1953.

McCracken, R. A. The informal inventory as a means of improving instruction. In Thomas C. Barrett (Ed.) The Evaluation of Children's Reading Achievement. IRA Perspectives in Reading, No. 8, Newark: International Reading Association, 1967, 79-97.

Smith, Frank. Understanding Reading. New York: Holt, Rinehart & Winston, 1971.

Weber, Rose-Marie. The study of oral reading errors: a survey of the literature. Reading Research Quarterly. Fall 1968, 4, 96-119.

Weber, Rose-Marie. A linguistic analysis of first-grade reading errors. Reading Research Quarterly, Spring 1970, 5, 427-451.

Hood

Table 1 Means, ranges and standard deviations of MISCUES made by REF and IMP second-graders

	Easier Story		More Difficult Story		Combined Story	
	REF	IMP	REF	IMP	REF	IMP
Mean	12.5	16.2	21.4	30.2	33.9	46.9
Range	4.4-57.6	2.8-31.8	3.2-71.8	13.2-60.2	9.8-129.4	16.0-92.0
Standard Deviation	11.25	6.60	17.66	14.22	27.92	18.93

Table 3 Mean proportions of corrections within miscue categories for REF and IMP second-graders and t-values for differences between means.

		Easier Story		More Difficult Story		Stories Combined		
		N	Mean	N	Mean	N	Mean	t-values
NCA-COR	REF	25	78.2	25	39.6	25	48.0	t=1.04 ($p \leq .302$)
	IMP	25	71.7	25	36.6	25	43.2	
PRE-COR	REF	23	72.3	24	41.8	25	52.5	t=3.10 ($p \leq .003$)
	IMP	25	59.9	25	23.6	25	36.3	
SENPASS-COR	REF	23	46.0	23	17.7	24	30.8	t=2.13 ($p \leq .038$)
	IMP	24	24.1	25	11.2	25	17.4	
GS-COR	REF	25	65.0	25	39.8	25	47.9	t=1.58 ($p \leq .120$)
	IMP	25	54.7	25	34.1	25	40.0	
NGS-COR	REF	25	59.4	23	35.6	25	46.4	t=3.24 ($p \leq .002$)
	IMP	24	40.5	25	24.0	25	30.6	

Table 2 Mean proportions of contextually appropriate and inappropriate, graphically similar, and corrected miscues for REF and IMP second-graders, t-values for differences between means, and reliabilities of mean scores.

		Easier Story Means	More Difficult Story Means	Stories Combined Means	t-values	R
NCA	REF	36.7	56.5	49.4	t=0.66 ($p \leq .515$)	.91
	IMP	31.8	54.9	47.3		
PRE	REF	28.3	29.3	29.3	t=0.82 ($p \leq .416$)	.88
	IMP	28.5	27.6	27.4		
SEN	REF	12.8	8.5	9.8	t=0.59 ($p \leq .557$)	.81
	IMP	14.4	8.6	10.7		
PASS	REF	22.1	5.6	11.7	t=1.24 ($p \leq .222$)	.94
	IMP	25.1	8.8	14.4		
SENPASS	REF	34.9	14.2	21.2	t=1.67 ($p \leq .102$)	.92
	IMP	39.6	17.4	25.2		
GS	REF	45.6	59.3	54.2	t=2.33 ($p \leq .024$)	.98
	IMP	36.6	49.9	44.9		
CORR	REF	62.4	36.7	45.9	t=3.12 ($p \leq .003$)	.98
	IMP	46.5	27.7	33.7		